CONTENTS

16.	Dr. Anatol W. Holt Applied Data Research, Inc.	NIC-8139
17.	Professor Leonard Kleinrock University of California, Los Angeles	NIC-8140
18.	Professor J.C.R. Licklider Massachusetts Institute of Technology	NIC-8141
19.	Dr. Thomas Marill Computer Corporation of America	NIC-8142
20.	Professor John McCarthy Stanford University	NIC-8143
21.	Mr. David McIntyre University of Illinois	NIC-8144
22.	Mr. Robert E. Millstein Applied Data Research, Inc.	NIC-8145
23.	Professor Marvin Minsky Massachusetts Institute of Technology	NIC-8146
24.	Professor Allen Newell Carnegie-Mellon University	NIC-8148
25.	Dr. Melvin Pirtle NASA Ames Research Center	NIC-8150
26.	Professor William K. Pratt University of Southern California	NIC-8151
27.	Dr. Bertram Raphael Stanford Research Institute	NIC-8152
28.	Professor J. A. Robinson Syracuse University	NIC-8153
29.	Professor Daniel Slotnick University of Illinois	NIC-8154
30.	Mr. Keith Uncapher The RAND Corporation	NIC-8156
31.	Dr. Richard Watson Stanford Research Institute	NIC-8158
32.	Dr. Richard Wishner Systems Control, Inc.	NIC-8157

THE ALOHA SYSTEM

(ADVANCED FORMS OF COMPUTER - COMMUNICATION NETWORKS)

The present generation of computer-communication systems is based on the use of leased or dial-up common carrier facilities, primarily wire connections. Under many conditions such communication facilities offer the best possible communications option to the overall system designer of a large computer-communication facility. In other circumstances however the organization of common carrier data communication systems seriously limits the possibilities of a large information processing system.

Since September 1968, THE ALOHA SYSTEM Project at the University of Hawaii has investigated alternatives to the use of conventional wire communications in a geographically diffuse computer system. When the constraint of data communications by wire is eliminated a number of options for different methods of organizing data communications within a computer-communications net are made available to the system designer. THE ALOHA SYSTEM Project has investigated and built a new and simple form of random access communications based on UHF radio links for a statewide university computing system.

The communications provided by THE ALOHA SYSTEM is capable of accommodating interactive alphanumeric consoles, graphic consoles, RJE terminals and minicomputers located within 100 to 200 miles of the central facility in Honolulu. Each remote location is linked to the central station by two 24,000 baud UHF links. Each location must format its information in a packet of up to 80 characters of text and 8 ID and control characters; the maximum packet length is thus 704 bits, requiring about 30 msec. to transmit. Thus each remote terminal must be interfaced into the system with a buffer/control unit connected to a modem and UHF radio transceiver.

The central station is connected to an IBM 360/65 with 2.5 Mbytes of core. The central station and the first remote terminal were put into operation in June 1971. At the present time three alphanumeric terminals are in operation, interfaced by means of buffer/control units designed and built by THE ALOHA SYSTEM. Efforts are now underway to expand the variety of equipment linked by the system. In particular the replacement of the hardware buffer/control unit by means of software packages in intelligent terminals, RJE terminals and minicomputers is in progress.

At the present time the software interface into the 360 is accomplished by means of a University of Hawaii time sharing system, UHTSS, built by THE ALOHA SYSTEM. On January 1, 1972, we plan to switch into IBM's TSO by means of an interface package recently completed and now under test.

The communication system now in operation is of greatest value in those situations where the ratio of the peak data rate to the average data rate is high, as in the interactive use of alphanumeric consoles. The present communication system has the capability of handling over 500 active alphanumeric consoles.

In order to provide the computing capabilities to match these communication capabilities the project has expanded its scope, as of September 1971, to include general studies of multiprocessor organizations. Such studies include specific treatment of the BCC 500 system, as THE ALOHA SYSTEM will acquire and make use of this equipment; the ILLIAC IV; and a number of other multiprocessor organizations being used or constructed under ARPA support. In January 1972, the BCC 500 equipment will be moved to Honolulu and will be set up to be operational within a calendar year. During this time research will begin on the following topics:

- analysis of the BCC 500 design
- a study of multiprocessing versus uniprocessing
- general methods for partitioning total system workload over a number of processors
- the use of a common memory system versus separate memory systems in a large multiprocessor system
- a study of the effects of long transmission delays in the design of a large system
- a study of the effects of geographical distribution of processing power on the performance and cost of a large system
- considerations of the impact of LSI technology on the design of special purpose processors
- analysis of ILLIAC IV and other specific multiprocessor organizations

Equipment will be designed and added to the BCC 500 system to assist in comprehensive measurements of the hardware and software under various conditions of load. Thus the system will be both an object of study and a tool for the study of others by providing extensive simulation capability. During this period it is planned to connect the 500 both to THE ALOHA SYSTEM communication network and to the ARPA network via satellite.

HARVARD CENTER FOR RESEARCH IN COMPUTING TECHNOLOGY

The ARPA Project in Calendar 1971

The Chemistry Project

LHASA (Logic and Heuristics Applied to Synthetic Analysis) is the name of an interactive Organic Chemistry program which has been gradually evolving over the past four or five years under the guidance of Professor E. J. Corey of the Chemistry Department. The approach chosen is to analyse the target compound, and generate a collection of molecules, each of which is one known chemical reaction away from the target. Then any of these is in turn analysed as the target, generating another collection of precursors. The analysis continues in the fashion in the "retrosynthetic" direction until one or more simple, readily available compounds result.

In order to codify the huge collection of known synthetic reactions it was necessary to classify them according to the type of molecular substructure resulting from the reactions. We have come up with five major classes, which we call (1) two-group transforms, (2) single-group transforms, (3) functional group interchanges, (4) functional group addition transforms, and (5) ring-oriented transforms. The first four classes are operative now in LHASA, and work is progressing on class (5).

Much of the work in the past in our project has centered around adding new reaction classes to the program. What we are beginning to implement now is a collection of synthetic strategies. These strategy modules will examine the structure and decide which collections of transforms would be best to apply, and how hard to work to clear the way for the application of certain powerful transforms. Look-ahead techniques will lead the analysis through seemingly counter-productive intermediate steps, if the end result is a drastic simplification of the structure. At all times of course, the chemist can retain full control of the direction of the analysis. Our philosophy so far has been toward the implementation of new strategies which will guide the search automatically, but still leave the chemist the option of manually selecting certain chemistry packages, as he has in the past. During the last year we have also done a good deal of work on teaching LHASA about stereochemistry.

With regard to the "intelligence" of the analyses produced by LHASA, we have processed a number of compounds that have appeared in the Synthetic Organic literature and have often come quite close to the published syntheses. In some cases LHASA has come up with more elegant solutions. While there are still many glaring holes in LHASA's chemical knowledge, its performance within its sphere of expertise has been very reassuring.

ECL Programming System and Laboratory for the Study of Automating Programming

The ECL programming system has been designed as a tool for tackling 'difficult' programming projects. Specifically, projects of this nature are systems characterized by two requirements: (1) Considerable experimentation is required to develop the system; that is, the design and development of the system must go hand in hand. (2) When a complete system is ultimately designed and programmed, it must be possible to take the working programs and produce a highly efficient product - both in machine time and space - with no change to the basic algorithms of their representation.

At the present time an experimental version of the system is operational - a version which only partially meets the above requirements. Additional system development is underway and will continue for some time.

The ECL programming system consists of a programming language, called ELl, and a system built around that language to provide a complete environment for the human-oriented use of the language. The system allows on-line conversational construction, testing, and running of programs. It includes an interpreter, a fully compatible compiler, and an editor - all callable at run-time, on programs constructible at run-time either by the programmer or as the result of computation.

One particular project employing ECL is currently underway at Harvard. That is the implementation of a laboratory for the study of programming automation techniques. Automating programming entails transfering to a computer those facets of programming which are not carried out efficiently by humans. It is our contention that the fact most in need of such a transfer is the optimization, in a very broad sense of the word, of programs.

As we see the problem of program automation, the issue is not arriving at a program, but arriving at a good one - one which operates with acceptable efficiency. Thus, programs produced mechanically (say by theorem provers applied to predicates), and even programs produced by a human programmer often leave much to be desired. The larger the program, the more likely this is to be the case. The reasons are generally such defects as inefficient representations of data, failure to notice or exploit possible constraints, use of inefficient or inappropriate control structure, redundant computations, inefficient search strategies, failure to exploit certain features of the intended host environment, etc. Recognizing the occurrence of these kinds of defects and remedying them is the primary goal of the laboratory.

The EL1 language plays three roles in the laboratory: (1) it is the language used to construct the various components of the system, (2) it and its extensions are the language used to state algorithms which are to be manipulated by the system, and (3) it is the target language for transformations (i.e. EL1 programs are transformed into better EL1 programs).

The features of EL1 and its host system, ECL, which are particularly relevant to the laboratory are the following: (a) Data types or "modes" can be programmer defined using several basic data types and recursively, several mode valued functions. (b) Procedures can be generic in the sense that a given procedure can have a number of different bodies or meanings and the selection of a particular body or meaning to be used is determined by the mode(s) of the argument(s) utilized in some call of the procedure. (c) ECL provides storage allocation and reclamation mechanisms which are quite sensitive to space/time efficiencies. There is a special compiler for data type definitions which goes to some lengths to utilize memory efficiently when allocating a component of a complex data structure containing several "pieces." The use of both "stack" and "heap" mechanisms for allocation and freeing of space is also provided. (d) ECL provides for multiple concurrent paths and permits complete user control over the environment for each concurrent path. (f) ECL includes primitives which facilitate writing "non-deterministic" algorithms. An EL1 program can make a tentative choice of one value from a set of possible values and proceed under the assumption that the chosen value is "correct," i.e. leads to a satisfactory result of the program. Should this assumption prove incorrect, the program can fall back to the point at which the choice was made. The entire environment of the choice point is restored and the program can proceed with a different value from the set.

The laboratory is essentially, one particular extension of ECL; the several EL1 programs which together with the ECL system will constitute the initial version of the laboratory are: a control program, a programmable theorem prover, a data base of transformations (theorems), a pattern matcher, and a package of measurement tools. At the present time we anticipate that an initial version of the laboratory will be operational during the summer of 72. At this time there will be a preliminary but useful set of "standard" transformations in the data base, plus a number of language extensions such as one hosting algorithm concerned with finite sets. Once the laboratory is operational we hope to have a population of users who are primarily interested in utilizing the facilities it provides in their development of non-trivial application programs. We also inted for the laboratory to provide a powerful research tool in the sense that workers in various areas will be able to implement their ideas and experiment with them at a cost which is drastically reduced from that which would be incurred if they had to "start from scratch."

Use of the ARPA Network

Harvard participated in the design of the HOST-HOST and file transfer protocols. The Network Control Program has been implemented on both the PDP-1 and PDP-10. In the former case, a new time-sharing monitor was designed and implemented (the old monitor was a single user monitor) in order to permit operation of the PDP-1 as a terminal processor. A single-user predecessor to the PDP-1 NCP which makes the machine a fancy terminal for the PDP-10 (via the network) has seen considerable use over the past six months; an updated version of this program which uses the NCP and new monitor will be operating soon.

Checkout of the PDP-10 NCP appears to be substantially complete. Starting late this week, we plan to have the NCP up five mornings a week. (The NCP takes 5K of core out of our current total of 31K user core. Many user programs cannot run in the reduced core space, hence our decision to run the NCP on a scheduled basis. A redesign of the NCP and IMP service routines will begin shortly. In the new version, we will run the NCP as a "user" program rather than its being a part of the resident monitor).

A file transfer program has been written and is approaching completion of its checkout phase. The program will work on both the standard DEC monitor and the TENEX monitor. The work on this program will serve as the basis for an extension to the ECL system which will allow transmission of data in internal ECL representations, accompanied by data descriptions. This will serve as the starting point for an investigation of techniques for transfer of data with change of representation (e.g., between PDP-10 and System/360 formats).

Howard Frank
Network Analysis Corporation

NIC 8134
12/8/71

RESEARCH IN STORE AND FORWARD COMPUTER NETWORKS

The present Network Analysis Corporation contract with the Advanced Research Projects Agency has the following objectives:

- To determine the most economical configurations for the ARPA Computer Network.
- To study the properties of store-and-forward networks, and in particular to investigate the relationship between traffic, routing, throughput, and cost.
- To develop procedures for the analysis and design of reliable and survivable computer and communication networks.

The research effort has resulted in a constantly evolving network optimization computer program which is able to produce extremely economical networks. The program's capabilities have been advanced to the point where networks with several hundred nodes can be handled. Cost-throughput characteristics for a 200-node store-and-forward network were determined. These characteristics extend the results of previous studies which showed that large ARPA-like networks are economical to operate using the present equipment of the ARPA net.

For the ARPA Computer Network, low cost networks have been derived and augmented as the network has grown. It has been shown that the ARPA network provides near optimal performance and retains its high throughput capabilities under variations in input traffic rates.

Activity in the reliability and survivability areas has focused on formulating realistic network survivability criteria and developing procedures for analyzing and designing large networks. The effort has resulted in the development of analysis methods more than 1000 times more efficient than conventional schemes.

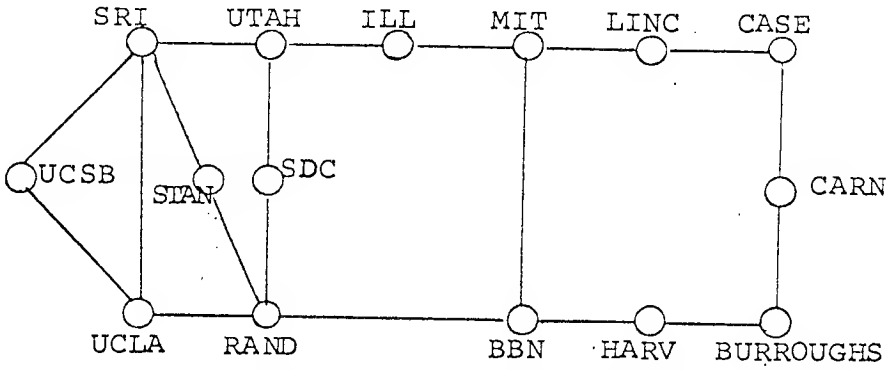
Present activities are centered in the following areas:

- Continue to develop the network optimization computer program for use in studying Defense Department computer and communication requirements.
- Develop the optimization program as a network resource for the ARPA Network.
- Provide design support to determine efficient network structures as additional nodes are added to the ARPA Network.

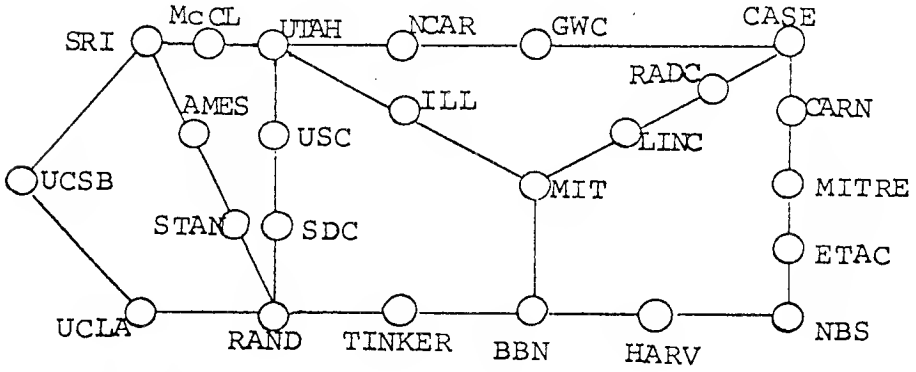
- Continue the study of the economics of large computer networks for resource sharing and communications. Investigate the use of new hardware and communication line options to enhance economy.

- Continue the study of reliability and survivability of computer and communication networks.

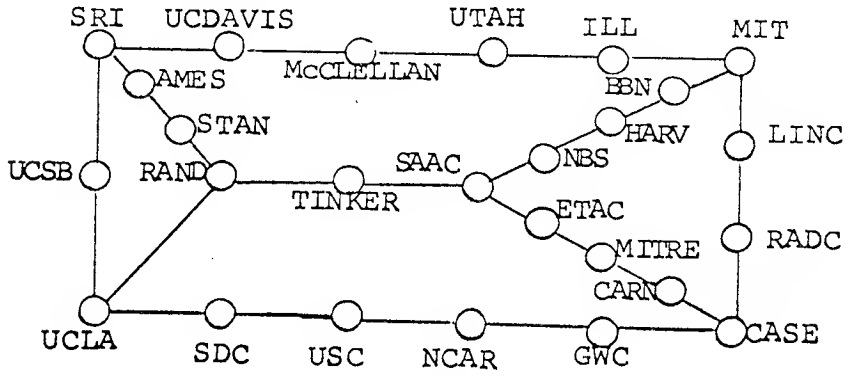
Shown below are several stages in the evolution of the ARPA Network design.



15-IMP NETWORK 3/1/71



24-IMP NETWORK 4/1/72



26-IMP NETWORK - PLANNED

UCSB COMPUTER RESEARCH LAB

I. ARPA Network

In this contract period, a Network Control Program was designed, implemented and modified as required to fully support the Host-Host protocol currently specified by the Network Working Group.

UCSB's NCP implementation has been thoroughly documented in a paper entitled "An NCP for the ARPA Network" (21 December 70, NIC 5480) and distributed to Network participants through the Network Information Center (NIC) at Stanford Research Institute in Menlo Park.

Access to the Network has been extended to include Fortran and PL/1 programs and On-Line System users, as well as assembly-language programs.

A subsystem of OLS called NET has been created to house, among other things, a set of operators providing access to basic, NCP functions from an OLS terminal. These operators were heavily used in early, interactive Network experimentation with (in particular) the Rand Corporation in Santa Monica. All of these operators have been described in detail in an RFC entitled "Network On-Line Operators" (21 April 71, NIC 5833) distributed to Network participants through the NIC. They are briefly described in the following paragraphs.

A User Telnet has been written in accordance with the protocol adopted to handle such teletype-like communication. The program conforms to the protocol adopted by the NWG to handle such communication. UCSB's User telnet implementation is described in detail in an RFC entitled "A User Telnet - Description of an Initial Implementation" (9 August 71, NIC 7176) distributed to Network participants through the NIC.

A number of services have been made available to Network users. Two separate specifications for access to the On-Line System have been designed and implemented, each supporting the graphic display features of the system. Both included provision for transmitting curvilinear display in line, dot, dot-dot, and character modes, and the first for transmitting special character display. The first such implementation was made available to Network users shortly after the adoption by the NWG of its initial Host-Host protocol specification, and was never used by any but local users. It was fully documented

in an RFC entitled "Specifications for Network Use of the UCSB On-Line System"(16 October 70, NIC 5417) distributed to Network participants. Support for this specification was terminated by UCSB this August, and replaced by generation of a new specification to which the Rand Corporation in Santa Monica has already interfaced its video-graphics system.

A remote job entry facility has been written, providing Network users with access to UCSB's batch processing facilities. Two independent processes were written to provide this service, each addressed by a standard ICP to a separate socket. One supports remote job entry (RJE) by accepting files of card images from the Network and transmitting them to UCSB's Houston Automatic Spooling Priority System (HASP) through an internal reader. The second retrieves output from remotely submitted jobs from a PDS and relays it to the Network user. These facilities have been in use on a production basis by user groups at the Rand Corporation in Santa Monica for a number of months, and are fully documented in an RFC entitled "Network Specifications for Remote Job Entry and Remote Job Output Retrieval at UCSB" (22 March 71, NIC 5775) distributed to Network participants through the NIC.

A Network file system has been written, making on-line, direct-access storage available to the Network community. This process (known as UCSB's Simple-Minded File System (SMFS)) has been fully documented in an RFC entitled "Network Specifications for UCSB's Simple-Minded File System" (26 April 71, NIC 5834) distributed to the Network community through the NIC.

Finally, a teletype-oriented interface to the On-Line System has been implemented according to the Telnet protocol adopted by the Network Working Group.

Maintenance and functional improvement of the UCSB IMP-HOST Interface has continued over the last year. Design improvements have been incorporated in a newly designed I.C. version of the interface for 360 operation. Several ARPA sites are interested in using the new interface, one of which is presently operational at MIT Lincoln Laboratory.

Hardware assistance in support of the UCSB On-Line System (OLS) has resulted in the development of a new Multi-Line Controller and a new version of the graphics display console used on the OLS. The Multi-Line Controller

(MLC) will allow the attachment of any type of user equipment to our 360 system.

The high-speed serial data link between the UCSB Network Host (IBM 360) and the speech analysis computer (SEL-810) has been designed and all purchasing has been completed. This link will be made operational in the next quarter, and will provide the hardware to link our speech system to the network. The speech software on the 810-B side of the 810B to 360 link has been completed and final checkout is underway. The speech software for the 360 side of the link should be completed during the next quarter. The 810-B to 360 link will allow up to two minutes of contiguous speech to be processed. In addition the 360 resident ASCØF speech software (analyzer, synthesizer, and filtering packages) will be available to network users.

II. Speech Project

A. Computer Classification and Recognition of Phonetic Information

The research program in the classification and recognition of phonetic information has solidly established that recognition parameters can be extracted from the ASCØF parameter set and used to perform reliable recognition of speech. This has been demonstrated by the following accomplishments:

1. Development of a segmentation procedure which reliably partitions the input speech data into basic voiced-unvoiced segments.

2. Utilization of the segmentation algorithm in the formulation of a single speaker vowel recognizer for recognizing steady state vowels and vowels embedded between two unvoiced phonemes using ASCØF data. The vowel recognizer is completely operational and achieved a recognition score of 97.4 percent correct identification for a data base of 240 recordings.

3. Extension of the segmentation algorithm for voiced-unvoiced phonemes to the detection of unvoiced phonemes.

4. Preliminary development of an unvoiced phoneme recognizer. A prototype unvoiced phoneme

recognizer for the 9 unvoiced phonemes has been implemented and is currently being tested. Although no recognition score has been obtained for the recognizer it is anticipated that it will be in the order of 90 percent accurate for the unvoiced phoneme set.

5. Extension of the segmentation algorithms of (1.) and (3.) to the segmentation of voiced phonemes. A segmenter which segments the incoming speech stream into voiced (both vowel-like and voiced consonants), unvoiced, and null (regions of no phonemic activity) phonemic segments has been implemented. The segmenter is currently being tested on a wide spectrum of words in order to measure its effectiveness and to modify it where necessary. A data base of 157 words has been created up to this point in time for this purpose.

   The segmenter represents a preliminary version of the final segmenter that will be used in a complete single speaker recognizer. The basic features of frequency and energy in each of the five frequency bands for the 157 words have been extracted and entered into appropriate files for the unvoiced and voiced segments of each word. These features will be utilized to structure the voiced consonant portion of the single speaker recognizer and to further validate the recognition accuracies of the already existing vowel and unvoiced consonant parts of the recognizer. It is anticipated that the recognizer should be completed by the end of this contract period and should be about 90 percent accurate for the set of 36 phonemes it was designed to recognize.

B. Data Compression

Intelligible data compressed ASCØF speech at 2500-3500 Baud has been obtained. This represents a reduction by a factor of 15 - 40 over the raw ASCØF data rate of 45,000 - 100,000 Baud for full words. Studies presently being conducted indicate a further reduction of 500 to 1500 Baud is feasible resulting in an overall data rate of 2000 Baud for data compressed ASCØF speech. Recognition of test words at these reduced data rates is quite good; however no formal testing has, to date, been done although procedures for the final testing of intelligibility are currently under consideration.

Bolt Beranek and Newman Inc. NIC #8138
ARPA Network Contract

Principal Investigators Conference
Washington, D.C. — 8 December 1971

A major project in the past year has been the final design, implementation and network installation of the Terminal IMP. This new machine is now operational at four network nodes and is performing essentially as predicted. In connection with the Terminal IMP we have:

- Shifted the emphasis of the development from the Honeywell 516 to the Honeywell 316, and accordingly struggled with the hardware problems of such a switch. Various issues of design imperfections, noise susceptibility, and improper cabling required seemingly endless attention.
- Completed the design of the Multi-Line Controller for the Terminal IMP, and elected to do the complete fabrication of this unit at BBN.
- Produced a substantial new terminal-handling program for the Terminal IMP, including the numerous functions necessary to act as a "mini-Host" and deal with an appropriate subset of Host protocol. In addition, several sizeable new test programs were needed and have been produced.
- Written and issued substantial new documentation regarding the Terminal IMP, including a complete revision of the *Specifications for the Interconnection of a Host and an IMP* (BBN Report No. 1822), and a new document called the *User's Guide to the Terminal IMP* (BBN Report No. 2183). A complete revision of the *Operating Manual* (BBN Report No. 1877) is being completed.
- Acquired and tested a variety of terminal devices with the Terminal IMP. These devices included: Infoton Vista 1-H, Imlac PDS-1 with hard copy unit, Execuport Model 300, Odec Model 1322 Line Printer, Inktronic Printer, IBM 2741, Calcomp Plotter, and various standard Teletypes. In addition, both card mounted and standard package modems have been tested with some of these terminal devices.

The Terminal IMP *includes a 316 IMP* which is functionally almost identical to the original 516 IMP. Stand-alone versions of 316 IMPs have been fabricated and tested with the network, and initial network installations are planned in the near future. In connection with two such planned installations, at Tinker and McClellan AFB, we undertook to provide "special Host interfaces" for Univac 418III computers. These interfaces have been fabricated and are being tested.

Additional 516 IMPs were installed at Illinois and Paoli, and both hardware and software retrofits were made at all network nodes to improve performance, repair troubles, and increase MTBF. The "maintenance" of the network, including the detective work required to localize and repair intermittent failures, has consumed considerable project energy. We are hoping that recent retrofits will ameliorate this problem. As an interesting detail, we retrofitted all IMPs to have a *wired-in* IMP number, allowing the software to become literally identical in all 516 IMPs, and providing an increased tolerance to certain kinds of failure.

Changes and improvements were made in the Network Control Center at BBN, including the design and implementation of a major new program. We are now using a 316 as a Host on the BBN IMP to monitor and summarize status reports and loading statistics from the network IMPs. We are keeping records of traffic at each node, both internode and intranode.

We have actively participated in the development of Host protocol and in various efforts to simplify use of the network by Host organizations:

- We contributed significantly to the design of the Telnet protocol, the file transfer protocol, the data transfer protocol, and others.
- We accepted the role of editor and publisher for "approved" protocols, and spent considerable effort generating a finished version of the Host/Host basic protocol.
- We compiled and published the *Resource Notebook*, an important aid to network use, and have been working with SRI on the maintenance of this document.
- We have been engaging in regular testing of Host availability, as a form of modest pressure towards the completion of unfinished network control programs at various nodes.

The basic IMP/IMP communication protocol still contains design imperfections which will affect performance under heavy loading. During the year, several sequential study efforts have been addressed to problems of congestion and routing in an attempt to improve the protocol before actual network loading rises enough to incur operational difficulty. As a partial improvement, we are now implementing a major program revision called "allocated RFNMs" and we are still studying aspects of the routing problem. We have gained considerable insight concerning the relation between packet queues in the IMP and performance.

As network growth proceeds, problems are arising which concern network access of some special kind. We have been engaged in several *ad hoc* studies of this sort:

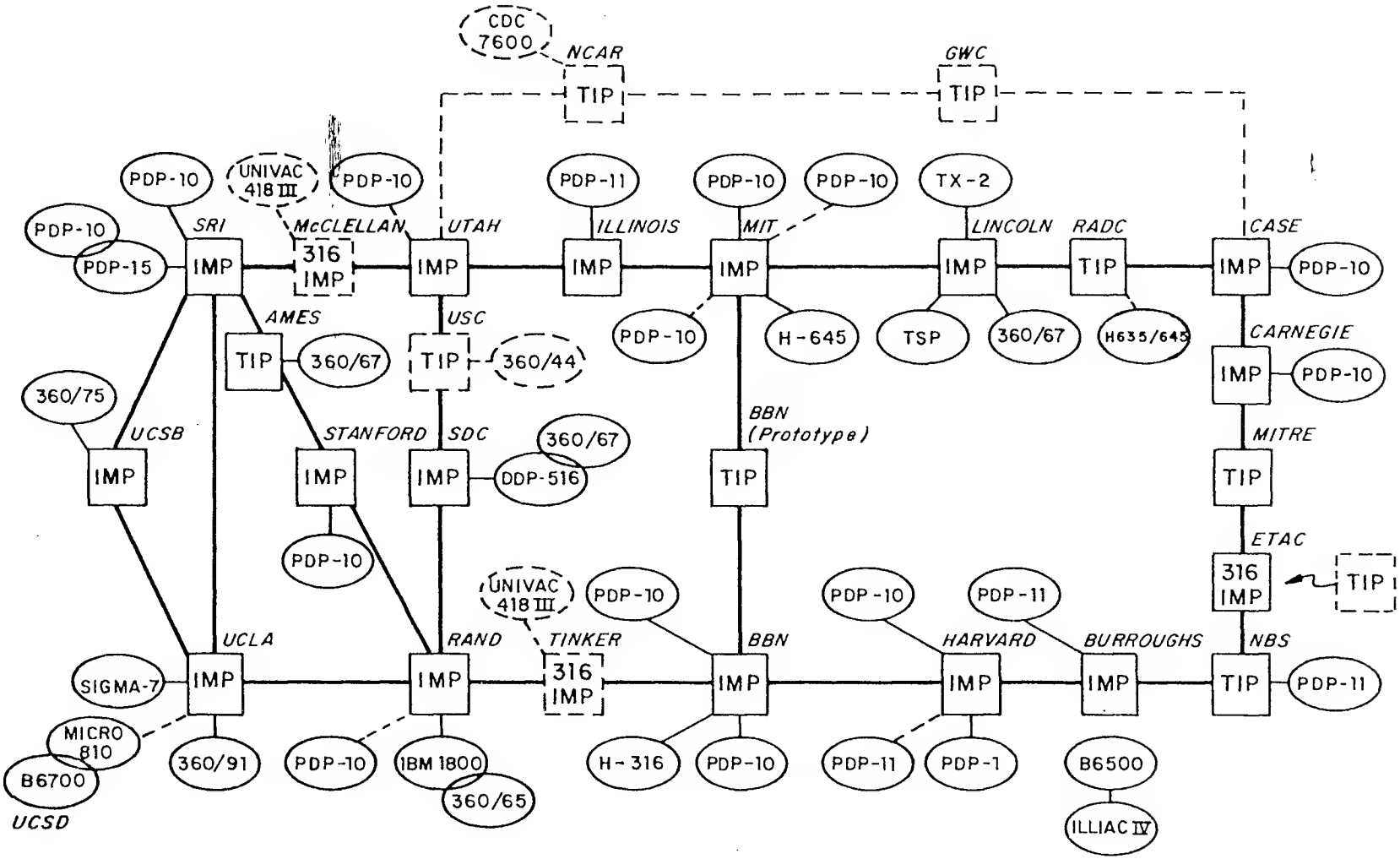
- We studied the problems of adding the Seismic Array Analysis Center to the net and the associated problems of sharing a transatlantic link between the seismic traffic (from Norway) and regular network traff
- We are studying the problems of "very distant Hosts" who wish to attach to an IMP from distances greater than 2000 feet.
- We are studying ways in which "remote job entry" type terminal configurations might sensibly gain access to the net.

As a major new initiative for next year, we have studied various possible designs for a "modular" IMP which could be configured to handle megabit bandwidth lines or alternately configured to provide a low cost version of an IMP. We are engaged in consideration of a mini-computer/multi-processor design which hopefully meets this need, but detailed design and choices of hardware have not yet been made.

The project continues to attract considerable attention and we have hosted numerous visitors and given numerous *ad hoc* presentations. In addition, a formal paper was presented at the "Second ACM Symposium on Problems in the Optimization of Data Communication Systems," and a formal paper on the Terminal IMP has been submitted for presentation at the 1972 SJCC.

Frank E. Heart
1 December 1971

IC #8138



ARPA NETWORK, DECEMBER 8, 1971

[DOTTED LINES INDICATE KNOWN PLANS]

UCLA SUMMARY REPORT

Leonard Kleinrock Principal Investigator
ARPA Contract No. DAHC-15-69-C-0285

At UCLA the principle activities associated with the ARPA Computer Network Research contract include modeling and analysis of computer systems (especially time-shared systems), modeling and analysis of computer networks, network measurements, and software development.

The common thread in virtually all of our research is the study of systems in which a collection of finite-capacity resources is subject to simultaneous demand; the resolution of this conflict and its effect upon throughput, cost and performance is our major objective.

1. Time-Shared Computer System Analysis

We are continuing our strong effort in studying the behavior of job flow and resource utilization for time-shared and multiprogrammed computer systems. We continue to study new classes of scheduling algorithms for single resource systems and have created large numbers of these algorithms with their associated analysis. We have also continued our effort in establishing general properties of all algorithms appropriate for CPU scheduling as reflected in our earlier results on precise bounds for response time functions. Currently, we are looking into the synthesis of optimum scheduling algorithms with respect to cost criteria which seem reasonable. We have now mounted a major attack on multiple resource scheduling problems so as to account not only for CPU allocation, but also for memory and other I/O contention. Here, too, we have been successful and can describe some of the important characteristics of these systems. We are currently extending our results to include non-exponential demands. An exciting new area is the application of the diffusion approximation in queueing theory to some of these multiple resource computer models. The diffusion approximation finds application in approximating computer network behavior as well. Along with these analytical studies, we have now begun to emphasize the validation of these models by the use of system measurement. We have carefully studied the UCLA scheduling algorithm and measurements indicate that significant improvements can be obtained (and some already have) by alterations to various of the scheduling algorithms and their parameters. Moreover, we are studying program behavior in a paging environment and are now investigating new resource allocation algorithms which make use of a priori information about a process.

2. Computer Network Analysis

Our major effort in modeling computer networks continues to move ahead. As these problems become more difficult, we find ourselves moving toward good heuristic network design, as opposed to optimal, idealized design methods. We are focusing more on the selection of capacities from a discrete set and are attempting to apply our former continuous capacity

optimization methods to this case as well. We are currently investigating a "flow deviation" method for assigning flows within a computer network; this is a suboptimal solution to the multicommodity flow problem which permits an iterative solution whose performance is rather good. We continue to use a detailed simulation of the ARPA Network to study routing procedures and network performance. Most of the routing procedures fall into the category of heuristic methods. This study has led us into the consideration of large network design; since the current routing tables grow linearly with the number of nodes in the network, it is clear that new procedures must be different in kind if they are to be useful for large networks. The obvious solution, long since proposed, is to create an appropriate partitioning of the network so that regional routing takes place at two or more levels.

3. Network Measurement

Now that the ARPA Network has a number of nodes with NCP's and now that the Network has grown and traffic activity is beginning to pick up, the network measurement center at UCLA is also becoming much more active. Our effort now is rather strong and we are creating specifications for a comprehensive measurement activity. In particular, we are investigating the properties of artificial traffic generators at remote Hosts, as well as at remote IMP's. We are also developing the notion of a traffic reflector which permits messages to be sent out over the net reflected back from some remote Host or IMP after being time-tagged and then sent on to some other destination (perhaps to be reflected again) and eventually recorded.

There are at least four motivations for carrying out these measurement studies. First, the obvious desire to measure the on-going activity and traffic within the network, as well as its performance and state. Secondly, to conduct experiments whereby we can measure the maximum throughput among various Hosts and IMP's within the network. Thirdly, to create experiments whereby we can measure the effect of planned changes on network performance; this is a rather difficult task and only some rather simple experiments are currently envisaged. Lastly, as an outgrowth of the measurement activity, we will be able to provide validation studies for our network models and analyses.

4. Software Development

In addition to keeping up with new protocol specifications we have made considerable effort to cooperate with other sites in the development of new network software. We have developed the software necessary to do remote job entry from our system to the IBM 360/91 at the UCLA Campus Computing Network. The Simple Minded File System at UCSB is now accessible at our site so that we can store and retrieve files at UCSB. We are currently completing the development of a compiler for the Data Reconfiguration Language which was developed at RAND. This will, for example, allow a user a convenient means of reformatting input data files so that they can be used as input for programs at remote sites on the network.

NIC 8142
December 3, 1971

Datacomputer Project

T. Marill
Computer Corporation of America

The goal of the datacomputer project is to develop a system capable of providing data-management and data-storage services having performance and cost characteristics suitable for on-line applications involving large shared files. When completed, the system is intended to serve as the major data handling facility in the ARPANET.

The system may be viewed as a black box having multiple physical ports at which external processors are interfaced. A standard notation, the datalanguage, defines a uniform mode of interaction between the external processors and the datacomputer. When data-language statements expressing service requests are presented at a datacomputer port, the system responds by executing the desired requests.

The types of service being planned are the following: (1) On-line storage of files and file descriptions, with an upper limit in the neighborhood of one trillion bits. (2) Retrieval of data, including whole files, subsets of files, and individual data elements; multi-key boolean retrievals will be executed without requiring sequential searches. (3) File maintenance functions, basically the addition of new data, deletion of old data, and changes to existing data. (4) Data reformatting, including conversion among character sets, number representations, and file formats. (5) Backup and recovery mechanisms, for use in case data is lost. (6) Accounting, for allocating charges to users. (7) Data security, preventing users from getting unauthorized access to data. (8) Data sharing, allowing the same data bases to be accessed by different users. (9) Simultaneous multi-user access, allowing multiple requests to be operated on simultaneously.

Current plans call for the datacomputer to have two physical ports, one being a wide-bandwidth channel to be interfaced to the Illiac IV system, the other a slower channel to be interfaced to an IMP. Computers in the ARPANET will thus have two alternate access paths into the datacomputer.

The major hardware components of the system consist of a modified PDP-10 computer, a bank of direct access storage devices, and a mass memory, currently planned to be a Precision Instrument Unicon 690.

During 1971 the system architecture, the file structures, access techniques and the datalanguage have been designed. Coding has begun.

Coordination has been set up with the following interested groups: RAND, University of Illinois, NASA/Ames, ETAC, Network Working Group. A new Working Group (Weather Data Base Working Group) has been established to deal with a particular database being implemented for the datacomputer. All potential datacomputer users have been polled and several have been interviewed to determine their requirements. Datacomputer design documents have been distributed through the NIC.

December 4, 1971

STATUS REPORT

ILLIAC IV Project Office

NASA/Ames Research Center

Mel Pirtle, Project Manager

During the last year, the ILLIAC IV Project Office has assumed a significant responsibility for the ILLIAC IV project, and will assume major responsibility during the upcoming calendar year, upon delivery of the equipment to Ames. To date, the ILLIAC IV Project Office has been involved primarily in administrative matters involved in establishing the Project Office and effecting the transition of responsibility from the University of Illinois. The Project Office reports to the Director of Research Support (Loren G. Bright) and has a staff of 10 people. It is supported by several contractors with expertise in systems design, hardware development, and management. It is anticipated that the Project Office staff will be increased to 20 during the next few months.

Very generally, the following goals have been established for the Project Office:

1. The ILLIAC IV complex with all of its resources shall be remotely accessible to all users via the ARPA Network, ultimately via several user interfaces appropriate for interactive, remote batch, graphic, and other modes of operation.

2. All owners of the ILLIAC IV complex shall have guaranteed resource allocations and hence shall be able to trade these among themselves. All users will relate to the system through an owner which will provide resource allocations and other services as appropriate.

To accomplish these goals, a design effort has been initiated on a remote access system in which to embed the ILLIAC IV. This system will include a central memory capable of serving as a buffer between the ILLIAC IV array, its disk, and other storage devices. This memory will also serve as the central memory for an augmented TENEX system. This system is being designed to optimize the utilization of the ILLIAC IV disk and to provide convenient methods of access to the system via the ARPA Network.

Also, the Computation Division at Ames has been assigned the responsibilities associated with NASA's ownership. Thus, this division, with Project Office support, is training users and providing access to ILLIAC IV simulators, etc. In particular, the GLYPNIR-to-PL1 translator is being run on the Computer Division's 360/67, and SSK is being run on the B6700 at the University of California at San Diego. Remote access to the latter facility is planned for early

next year.

The current status of some of the hardware and software is as follows:

1. The ILLIAC IV hardware has been fabricated and is undergoing checkout. Burroughs' estimates for completion of checkout vary from March to July -- all unrealistically optimistic.

2. The software developed at the University of Illinois, including the operating system OS1, the languages GLYPNIR and ASK, and the simulator SSK are operational and have been delivered to the Project Office.

3. A leased PDP-10 system has been installed at Ames and is running the ALCOM system.

4. The UNICON ($.7 \times 10^{12}$ bit laser memory) has been delivered and connected to the PDP-10. Final acceptance tests will begin next week.

During the next several months, an interim system will be developed which will provide for final checkout and early use of the ILLIAC IV. This system will consist of the PDP-10 (with TENEX operating system) and possibly the UNICON, connected to the ARPA Network and to the ILLIAC IV system via its BIOM (Buffer I/O Memory). Thus, the PDP-10 will connect to the ILLIAC IV in parallel with the B6500. This system will be used to augment the test and exercise programs now available on the B6500 and will provide the simplest, most reliable operating system for the ILLIAC IV that can be reasonably developed.

Summary of Activities of Center for Advanced Computation during 1971

Information Management System

- Work continued on the design of an Information Management System for ILLIAC IV. An investigation of the capabilities of ILLIAC IV to perform data management functions concluded that ILLIAC IV may be cost-effectively employed for general sequential-access data bases and for special well-structured random-access data bases.

ARPA Network

- Completed Version I of the ARPA Network Terminal System (ANTS) based on a "mini-host PDP-11 computer. This system allows sophisticated connection of terminal devices to the network and provides local housekeeping services to the network site.

- A basic graphics display package for use on the B6700 was completed. Displays were integrated into ANTS and a direct connection placed between the B6700 and the PDP-11 computer providing graphical output capability to the B6700.

- The systems implementation language for the PDP-11 was transferred to the B6700. A complete programming language manual was written.

- A document called the "Network User's Handbook" was created to provide a scheme for coordinated and coherent distribution of documentation specifically pertaining to the ARPA network and its usage by Center and University personnel. The notebook covers the major areas of CAC network activities and systems, the specific service sites and their usage on the network, the Network Information Center (NIC) and its access, and compendiums of information concerning the various research sites, their activities, and their usage of the network.

Numerical Analysis

- The following ILLIAC IV programs have been fully debugged and tested on the simulator:

  1. The Algebraic Eigenvalue Problem

     - A Jacobi algorithm for Hermitian matrices with a high rate of convergence (ASK).

     - Finding all of the eigenvalues and eigenvectors of a symmetric tridiagonal matrix for the core contained case ($n \leq 192$).

2. Solution of Linear Systems of Equations

- Gaussian-Elimination, $n \leq 1000$ (ASK).

- The method of modification, $n \leq 64$ (ASK).

3. Solution of Nonlinear Equations

- The Secant method for finding a root of the single nonlinear equation $x = f(x)$, (ASK).

- An algorithm for finding real roots of real polynomials (ASK).

4. Time Series Analysis

- Computing the zero-phase impulse response of a band-pass filter, (GLYPNIR).

- Fast Fourier transform of a real function, (GLYPNIR).

- Sorting of a set of points in three-dimensional space in ascending order with respect to their distances from a point on the z-axis, (GLYPNIR).

- A Fast Fourier transform procedure for the ILLIAC IV (ASK).

5. Ordinary and Partial Differential Equations

- A Block-Iterative method for solving a second order elliptic partial differential equation on the ILLIAC IV, (GLYPNIR).

- The ADI method for solving a second order partial differential equation on the ILLIAC IV, (GLYPNIR).

Education

- Education of segments of the ILLIAC IV user community was accomplished through seminars, classes, and the development and dissemination of tutorial materials.

- A document entitled "An Introductory Description of the ILLIAC IV System" was completed for use by application programmers who would like a tutorial description of the ILLIAC IV system before attempting to read the Systems Characteristics and Programming Manual.

THE RAND CORPORATION
1700 Main Street
Santa Monica, California 90406

OVERVIEW OF RAND'S IPT RESEARCH IN MAN-MACHINE SYSTEMS

RAND has assumed some interesting and important new tasks this contract period. These tasks as reported here, along with on-going projects, are focused on the application of advanced computer-based techniques to new areas of concern in the national defense posture. Reported here also is the development of tools to support these projects and provide new resources for Network users.

ADVANCED AUTOMATION STUDY (T.O. Ellis, R.H. Anderson)

We are conducting a study of computer-based advanced automation of manufacturing processes, and the potential impact such automation could have on U.S. defense production and preparedness.

To focus this study, we are now examining the state-of-the-art of fabrication and assembly of rather small electro-mechanical products, and have inspected a number of relevant manufacturing facilities. In order to evaluate the potential impact of computer-based automation and to exemplify those areas of the technology which need additional development, we intend to develop specifications for a hypothetical manufacturing facility which is capable of highly automated yet flexible production over a range of products. This facility will incorporate the applicable advanced techniques in information acquisition and management, A.I., robotics, sensors, computer-aided design, operating systems, and adaptive machining. This hypothetical facility will then be compared to current production techniques to evaluate potential payoffs. Some impact areas we intend to investigate are: (1) shorter design and production cycles for "long lead-time" items, (2) easier design experimentation for prototypes, (3) formalization of product descriptions, thereby increasing product manufacturability and maintainability, (4) improved controls and better understanding of costs and schedules, and (5) increased productivity and cost effectiveness.

COMPUTER SECURITY ASSURANCE (R.M. Fredrickson)

Resource-sharing computer systems with remotely-located access facilities promise major benefits to military and intelligence communities if associated security problems can be solved.

A key problem in the application of resource-sharing systems within the classified information processing context is accurate assessment of system weakness due to improper design and/or implementation of operating system software. This project is directed to the development of techniques and tools of software analysis which can, if possible, establish that the safeguards designed into the software of a system are operative, function as intended, and collectively constitute acceptable controls for safeguarding classified information.

The project is approaching its task in two time frames. The shorter is concerned with the development of audit aids, information and flow chart systems and similar analysis tools which can be of immediate assistance to those faced with the responsibility of judging the security characteristics of a system. As a longer range goal, the project is attempting to specify the generic characteristics of operating system security and the system assurance process as it relates to security so that a more precise technology of system evaluation will be possible.

MICROPROGRAMMING PROCESSOR DEVELOPMENT (T.O. Ellis, M.R. Davis)

The Standard Computer Corporation MLP-900 is a prototype microprogrammed central processor with a very flexible programming structure.

The microprogramming processor is intended to be molded specifically to application areas or high-level language interpretation. Being writable, this context will be exchangeable, defining different machines at different times, in a manner analogous to the task switching which occurs among users of conventional time-sharing and multitasking systems. Capabilities for high-speed swap of context and microprogram are being included to enhance the time-shared use of this facility under control of the ARPA PDP-10 at RAND.

Planned as a major component of the PRIM system, the MLP-900 was only partially complete at the time Standard elected to stop development and donated the prototype to the project. The project's objective is to complete the design, fabrication, and check-out of the MLP-900, and interface it with the PDP-10 so that it becomes a viable part of the PRIM system.

The target completion date for the hardware implementation is November 1972.

PROGRAMMING RESEARCH INSTRUMENT (PRIM) (R.M. Fredrickson)

The project encompasses the installation, operation, and development of the PDP-10/MLP-900 system acquired to support IPT research programs at Rand and similar research programs across the country through the ARPANET.

As an interim goal, the PDP-10 is being prepared to provide PDP-10 TENEX service to both RAND and the Network-based ARPA-IPT programs. Combining the PDP-10 with the MLP-900 as a dual processor system providing multiuser access to microprogramming capability has been deferred, pending completion of the MLP-900 hardware (approximately November 1972). Regular TENEX service to the ARPANET community is scheduled to begin in January 1972.

Another component of the system is the Standard Computer Corporation's IC-4000. This stand-along computing system will be prepared as a remote job entry station to the PDP-10 to serve the special requirements of the ARPA-IPT sponsored RAND Climate Dynamics project.

Beyond this and the continuing development of Network capability (see ARPA Experimental Computer Network), we do not expect further work on the services offered by this facility until the MLP-900 becomes available toward the end of this contract period.

ARPA EXPERIMENTAL COMPUTER NETWORK (J.F. Heafner, E.F. Harslem)

RAND's participation in the ARPANET is part of a larger program in collaboration with other ARPA-sponsored researchers to distribute resources not otherwise accessible. Our goals are to provide local access to remote programs and facilities, to provide remote access to our local services, and to explore new techniques of ARPANET communications.

We have provided local access to remote services by developing and offering a Network Service Program (NSP). The NSP offers the user high-level command languages to such diverse services as the Remote Job System at UCLA and the On-Line System at UCSB. The NSP is used for daily production work by the Dynamics of Climate project at RAND. Use of NSP and the ARPANET has greatly increased the climate researchers' throughput over their previous mode of operation.

We have recently completed a graphics package which allows the user to graphically configure the tabular information (now received over the ARPANET) as contours, graphs, lists, and vector plots on our video consoles.

RAND contributes greatly to the emergence of an ARPANET community project, i.e., an experimental data reconfiguration service. The service provides one solution to the problem of communicating between two computer processes that have different input/output requirements.

ADAPTIVE COMMUNICATOR (R.H. Anderson, W.L. Sibley)

The aim of the Adaptive Communicator (A/C) Project is to make the man-machine interface more flexible and man-oriented. The A/C approach is based on three design principles: (1) that the communication task should be separated from the application program; (2) that multiple communication channels should be used simultaneously and interchangeably; (3) that the communication agent should be adaptive (i.e., trainable).

The project's major effort has been the design and testing of a notation and an operational scheme that will support the above design principles. In addition to these principles, the notation must allow for the priming of the agent with a priori knowledge. The notation is in the form of sets of ordered pattern-replacement rules; both pattern and replacement parts are expressed as directed, labeled graphs (webs).

A complete description of the notation and the operational scheme is given in R-876-ARPA, "Heuristic Context Analysis: The Basis for an Adaptive Man-Machine Interface."

COMPUTER-ASSISTED SPECIFICATION OF ALGORITHMIC PROCESSES (R.M. Balzer)

Due to the Project Leader's leave of absence, minimal progress has been made on CASAP since the last report. Work will continue on this project as time permits from the duties involving a new ARPA task in Automatic Programming currently being formulated.

RWW 7 DEC 71 0150

R.W. Watson
Network Information Center
SRI-ARC

Summary of 1971 Activities 1

The major accomplishment of the Network Information Center (NIC) during 1971 is that an evolving center came into operation offering both online and offline services to serve the ARPA Network Community. 2

The services presently offered by the NIC are a subset of capabilities being developed within the SRI Augmentation Research Center (see NIC--8129,) and have as their basic objectives: 3

(1) To help members of the Network Community find the people, system, or information associated with work going on over the Network that can help them fill their needs. 3a

(2) To help a geographically distributed group collaborate. 3b

The initial NIC services now available to meet the above goals are the following: 4

Online: 4a

(1) Access to the typewriter version of the Augmentation Research Center Online System (NLS) for communique creation, access, linking, and for experimental use for any other information storage and manipulation purpose suitable for NLS and useful to Network participants. 4a1

(2) Access to Journal, Number, and Identification Systems which allow messages and documents to be transmitted to Network participants. 4a2

(a) Documents or messages entered in the Journal System are maintained online for later viewing, using the facilities of NLS. 4a2a

(b) Document Distribution: 4a2b

Documents are now distributed by 4a2b1

i) placing the message or a link to the document in the receiver's "ident file". 4a2b1a

ii) in hardcopy form through the U. S. Mail. 4a2b1b

Documents will shortly be distributed through the Network when sites have implemented the File and Mailbox Transfer Protocols. 4a2b2

(c) A unique number is assigned each entry at the time of submission. Number(s) can also be preassigned for allowing related documents to be interlinked at the time of their preparation. 4a2c

(d) A catalog entry is prepared at the time of submission and later this entry is used to update a catalog kept both online and in hardcopy form. 4a2d

(e) Special interest groups can be created to facilitate indicating to the system distribution lists for dialogue items. Dialogue items can be placed in subcollections associated with the dialogue groups for special index production. 4a2e

(3) Access to a number of online Functional Documents through a special Locator file using NLS link mechanisms. 4a3

(a) The NIC collection catalog 4a3a

(b) ARPA Network Resource Notebook 4a3b

(c) NIC user documentation 4a3c

(d) Directory of Network Participants 4a3d

(e) Network Protocols 4a3e

(f) Links to other files created by sites with information of potential Network-wide interest. 4a3f

Offline 4b

(1) A Network Information Center Station set up at each site with: 4b1

(a) A Station Agent to aid use of the NIC. 4b1a

(b) A Liaison to provide technical information about his site. 4b1b

(c) A Station Collection containing a subcollection of documents of interest to Network participants. 4b1c

(2) Techniques for maintaining Functional Documents, such as: 4b2

(a) NIC Catalog 4b2a

(b) ARPA Network Resource Notebook 4b2b

(c) Directory of Network Participants 4b2c

(d) NIC User Documentation or Site Documentation 4b2d

(e) Other 4b2e

(3) Support of Network dialogue existing in hardcopy form through duplication, distribution, and cataloging. 4b3

(4) General Network referral and handling of document requests. 4b4

(5) Building of a collection of documents potentially valuable to the Network Community. Initial concentration has been on obtaining documents of possible value to the Network builders. 4b5

(6) Crude selective distribution to Station Collections. 4b6

(7) Training in use of NIC services and facilities. 4b7

We urge individuals and special interest groups within the Network Community to avail themselves of the above services and to let us know where we can improve existing services and what new services we should be providing. 5